

# Weighted distances between preferences

Citation for published version (APA):

Can, B. (2012). *Weighted distances between preferences*. METEOR, Maastricht University School of Business and Economics. METEOR Research Memorandum No. 056  
<https://doi.org/10.26481/umamet.2012056>

**Document status and date:**

Published: 01/01/2012

**DOI:**

[10.26481/umamet.2012056](https://doi.org/10.26481/umamet.2012056)

**Document Version:**

Publisher's PDF, also known as Version of record

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

Burak Can

**Weighted Distances Between  
Preferences**

RM/12/056

**METEOR**

Maastricht University School of Business and Economics  
Maastricht Research School of Economics  
of Technology and Organization

P.O. Box 616  
NL - 6200 MD Maastricht  
The Netherlands

# Weighted Distances Between Preferences

Burak Can\*

November 2012

## Abstract

Individual disagreements are assumed to be reflected in the preferences. Distance functions, e.g., the well-known Kemeny (1959) distance, are used to measure these disagreements. However, a disagreement on how to rank the top two alternatives may be perceived more (or less) than a disagreement on how to rank the bottom two alternatives. We propose two conditions on distance functions which characterize a class of weighted distance functions. This class allows to quantify disagreements according to where they occur in preferences. We provide some examples within this class and show one of them to be the generalization of the Kemeny distance on strict preferences.

*JEL classification:* C63, D71, D72, D74

*Keywords:* Strict preferences; rankings; Kemeny distance; Ideological distance

## 1 Introduction

In various contexts, the analysis of differences or dissimilarities between those opinions are crucial. Consider, for instance, a situation where like-minded people form clusters, interest groups, or political parties to implement their agenda on some institution. Another example would be situations in which dissimilarities between the social preference and that of the individuals cause discontent. In that case the extent of the discontent is very much dependent on the model of dissimilarity. It is important, thus, to measure how similar (or dissimilar) two individuals are regarding their preferences.

Preferences are often modeled as orders/rankings over available alternatives. To compare two preferences, it is, therefore, plausible to look at the alternatives which are ranked

---

\*Department of Economics, School of Business and Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands; E-mail: [b.can@maastrichtuniversity.nl](mailto:b.can@maastrichtuniversity.nl), Phone number: +31 43 3883807. I would like to thank the Netherlands Organisation for Scientific Research (NWO) for the grant with project nr. 400-09-354. I also thank Norman Schofield and Uğur Özdemir for their comments on earlier drafts while I was a guest at Weidenbaum Center on the Economy, Government, and Public Policy in Spring 2011.

oppositely. The well-known Kemeny distance (Kemeny, 1959) is commonly used in that way. Consider a strict preference  $a \succ b \succ c$ , which is interpreted as:  $a$  is preferred to  $b$ , and  $b$  to  $c$ , and by transitivity,  $a$  to  $c$ . The Kemeny distance between  $a \succ b \succ c$  and another strict preference  $b \succ a \succ c$  is 1, because the two preferences only disagree on how to order  $a$  and  $b$ . However, the Kemeny distance between  $a \succ b \succ c$  and  $a \succ c \succ b$  is also 1 (the disagreement is now on how to order  $b$  and  $c$ ). Therefore, the dissimilarity between the former two and the dissimilarity between the latter two are given identical weights. It is not that difficult though to imagine situations where a disagreement at the top of a preference leads to a larger conflict/dissimilarity than a disagreement at the bottom of the preferences.

We believe the variation on the dissimilarities caused by the position of disagreements in preferences might be useful in many applications. For instance, consider three search engines, *(G)oogle*, *(Y)ahoo* and *(B)ing*. Given a word search, assume these engines give a strict ranking of the same millions of alternatives, i.e., websites that are relevant to the search term. Suppose that  $G$  and  $Y$  provide identical rankings in the first hundred websites and differ completely in the remaining millions of websites. Suppose also that  $G$  differs from  $B$  in the ranking of the first hundred websites but is identical in the remaining millions of websites. Nevertheless, it is natural to argue that  $G$  is closer to  $Y$  than it is to  $B$ , even if  $G$  and  $Y$  disagree on how to rank the remaining millions of alternatives after the first hundred websites. This is because what apparently matters most for internet users in website rankings is the first twenty-thirty websites (BBC<sup>1</sup>, 2006) that are ranked. Another branch of applications would occur in cases where at least two individuals need to find consensus by making concessions, such as in bargaining or collective decision making. The implicit cost of these concessions, then, might depend on the positions of the disagreement between the individuals.

Note that the variation in dissimilarity may not always follow a monotonically decreasing pattern in the position of disagreement. In fact, in cases where certain positions in preferences are critical, the dissimilarity caused by a change in those positions might be more than changes in other positions. An example would be the ranking of football teams in a league, where the last, say 3, teams of the last week's ranking are to be relegated to another league. Then, a swap in the last 3<sup>rd</sup> and 4<sup>th</sup> positions might be much more critical, hence influential in the dissimilarity between two rankings, than a swap elsewhere. Therefore, it makes sense to assign more dissimilarity to a change at those critical positions.

In this paper, to model dissimilarity between preferences, we propose to use certain functions on strict preferences<sup>2</sup> in a similar spirit of the Kemeny distance, i.e., respectful to the number of disagreements, but also allow variation in the treatment of different pairs of disagreements. To that end, we distinguish between metrics (distance functions) and semimetrics (or dissimilarity functions) which are -roughly speaking- metrics that do not

---

<sup>1</sup><http://news.bbc.co.uk/2/hi/technology/4900742.stm>.

<sup>2</sup>A strict preference on a set of alternatives is a complete, transitive and antisymmetric binary relation over that set of alternatives.

necessarily satisfy the triangular inequality condition. This allows for more functions to analyze preferential differences. Although triangular inequality condition is very relevant, even for the very space we live within, when it comes down to the likelihood of preferences it may not always be that realistic. For instance, consider three parties (l)eft, (r)ight and (c)entre and three voters with preferences respectively:  $l \succ c \succ r$ ,  $r \succ c \succ l$ , and  $c \succ r \succ l$ . Note that voters have single-peaked preferences on the left-right political dimension. Obviously one can stereotype the first individual as a leftist, the second as rightist and the third somewhat centre-right. In case, the ideological dissimilarity between the leftist and rightist individuals is relatively high, modeling the dissimilarity with the triangular inequality condition may be difficult. It is, nevertheless, a likely scenario that people may underrate their political differences with middle individuals whereas they overrate them with individuals in completely opposing ideologies.

In this paper, we provide two conditions that essentially characterize a class of dissimilarity functions, which we call the “weighted dissimilarity functions”. First one, “positional neutrality” is a neutrality condition towards the position of disagreement between two *adjacent* preferences, i.e., preferences which have only one disagreement. The second one, “decomposability” is a additivity-like condition which requires that the dissimilarity between any two preferences can be decomposed into a *path* of adjacent preferences between the two. We show some examples from the class of weighted dissimilarity functions, hence semimetrics, and also prove under which conditions these functions become metrics, i.e., they satisfy the triangular inequality condition. We restrict our attention to strict preferences only and employ some group-theoretic results.

In Section 2, we introduce the notation and basic conditions for dissimilarity functions over strict preferences and define the two aforementioned conditions; positional neutrality and decomposability. In Section 3 we introduce the class of weighted dissimilarity functions and discuss some members of this class: the Kemeny distance, the Lehmer function, the inverse Lehmer function, and the path-minimizing function. Section 4 shows under what type of weight distributions, these functions become weighted distances, i.e., they satisfy the triangular inequality condition. It is shown that only for one weighted distance function, the path-minimizing distance, the variation in the weights does not affect the triangular inequality condition. However for the other two functions particular weight distributions guarantee the triangular inequality condition. Section 5 concludes the paper and points to possibilities for further research.

## 2 The Model

### 2.1 Notation

Let  $A$  be the set of alternatives with cardinality  $m \geq 3$ . Strict preferences are modeled by linear orders<sup>3</sup> over  $A$ , and the set of all linear orders is denoted by  $\mathcal{L}$ . Given  $R \in \mathcal{L}$ ,

---

<sup>3</sup>Complete, transitive and antisymmetric binary relations.

$a R b$  is interpreted as  $a$  is strictly preferred to  $b$ , i.e., the ordered pair  $(a, b) \in R$ . We sometimes write  $R = \dots a \dots b \dots$  if  $a R b$ , and  $R = \dots ab \dots$  if  $a R b$  and there exists no  $c \in A \setminus \{a, b\}$  such that  $a R c$  and  $c R b$ , i.e.,  $a$  and  $b$  are adjacent in  $R$ . Given any  $a \in A$ ,  $UC(a, R) = \{b \in A \mid b R a\}$  is the “upper contour set” of  $a$  in  $R$ , i.e., the set of alternatives that are ranked above  $a$  in the linear order  $R$ . Correspondingly,  $LC(a, R) = \{b \in A \mid a R b\}$  is the “lower contour set” of  $a$  in  $R$ .

For  $l = 1, 2, \dots, m$ ,  $R(l)$  denotes the alternative in the  $l^{th}$  position in  $R$ , and we use  $rank(a, R)$  to denote the position of alternative  $a$  in  $R$ . For some subset  $B \subseteq A$ ,  $R|_B$  denotes the preference reduced to  $B$ , i.e.,  $R|_B = R \cap (B \times B)$ . Given any two linear orders  $R, R' \in \mathcal{L}$ , the set difference  $R \setminus R'$  denotes the set of ordered pairs that exist in  $R$  and not in  $R'$ , i.e.,  $\{(x, y) \in A \times A \mid x R y \text{ and } y R' x\}$ . Two linear orders  $(R, R') \in \mathcal{L}^2$  form an *elementary change*<sup>4</sup> in position  $k$  whenever  $R(k) = R'(k+1)$ ,  $R'(k) = R(k+1)$  and for all  $t \notin \{k, k+1\}$ ,  $R(t) = R'(t)$ , i.e.  $|R \setminus R'| = 1$ . Given any two distinct linear orders  $R, R' \in \mathcal{L}$ , a vector of linear orders  $\rho = (R_0, R_1, \dots, R_k)$  is called a *path* between  $R$  and  $R'$  if  $k = |R \setminus R'|$ ,  $R_0 = R$ ,  $R_k = R'$  and for all  $i = 1, 2, \dots, k$ ,  $(R_{i-1}, R_i)$  forms an elementary change. For the special case where  $R = R'$ , we denote the unique path as  $\rho = (R, R)$ .

A bijection  $\pi : \{1, 2, \dots, m\} \rightarrow \{1, 2, \dots, m\}$  is called a *permutation* and the set of all permutations is denoted by  $\Pi$ . We use  $\pi(R)$  (or  $\pi \cdot R$ ) to denote the permutation of the linear order  $R$  by  $\pi$ , i.e.,  $\pi(R) = R'$  if and only if  $R(i) = R'(\pi(i))$  for all  $i = 1, 2, \dots, m$ . Given  $R, R' \in \mathcal{L}$ , a permutation  $\pi \in \Pi$  is called the *corresponding permutation*<sup>5</sup> for  $R, R'$ , if  $\pi(R) = R'$ . We denote the *conjugate* of a permutation  $\pi$  by  $\tilde{\pi} \in \Pi$ , i.e.,  $\tilde{\pi}(R') = R$  if and only if  $\pi(R) = R'$ . A permutation that swaps the  $k^{th}$  alternative of a linear order with  $(k+1)^{th}$  is called an *elementary permutation* and is denoted by  $\sigma_k$ . Hence,  $\sigma_k$  is the corresponding permutation for any  $R, R' \in \mathcal{L}$  that form an elementary change in position  $k$ . The set of all elementary permutations is denoted by  $S = \{\sigma_1, \sigma_2, \dots, \sigma_{m-1}\} \subseteq \Pi$ . The identity permutation is denoted by  $\sigma_0$ .

Note that the set of all permutations  $\Pi$  over the set of alternatives  $A$  forms a symmetric group (also known as a permutation group) with the group operator “ $\cdot$ ”, which implies any permutation  $\pi \in \Pi$  can be obtained by composition of some other permutations with the group operator, e.g.,  $\pi'' \cdot \pi' \cdot R = \pi \cdot R$  refers to the situation where  $R$  is first permuted via  $\pi'$  and then  $\pi''$ , and  $\pi'' \cdot \pi' = \pi$ . Note, however, that unless  $m \leq 2$ , the group fails commutativity, e.g., for  $R = abc$ ; note that  $\sigma_1 \cdot \sigma_2 \cdot R = cab$  whereas  $\sigma_2 \cdot \sigma_1 \cdot R = bca$ .

In this paper, we will especially make use of compositions of elementary permutations in  $S$ . Since  $\Pi$  is a permutation group it has  $S$ , as the generator set, which means every permutation  $\pi \in \Pi$ , including the identity permutation  $\sigma_0$ , can be expressed by some composition of elements of  $S$ . Given  $R, R' \in \mathcal{L}$ , and a corresponding permutation  $\pi \in \Pi$ , let  $I(\pi)$  denote the size of  $\pi$ , which is the number of minimal inversions required to obtain  $R'$  from  $R$  by elementary permutations. Note that as  $\pi$  is the corresponding permutation for  $R, R'$ , we have that  $I(\pi) = |R \setminus R'|$ . Note also that for the identity permutation, we

<sup>4</sup>We omit the paranthesis whenever it is clear and write  $R, R'$  instead.

<sup>5</sup>We omit this expression whenever it is clear which permutation we employ.

have  $I(\sigma_0) = 0$ . Next we define compositions of a permutations via elementary/identity permutations.

**Definition 1. Factorization:** Given  $\pi \in \Pi$  and some positive integer  $r \geq I(\pi)$ , a vector of elementary/identity permutations  $f = (f(1), f(2), \dots, f(r)) \in (S \cup \{\sigma_0\})^r$  is called a *factorization* of  $\pi$  whenever for all  $i = 1, 2, \dots, r-1$ :

- a)  $f(r) \cdot f(r-1) \cdot \dots \cdot f(1) = \pi$ ,
- b)  $f(i) \neq f(i+1)$ .

If  $r > I(\pi)$ , i.e., the number of inversions is not minimal in the factorization  $f$ , then  $f$  is called a *non-reduced factorization* of  $\pi$ . Next we define the minimal factorizations, i.e., the compositions that require the fewest elementary permutations.

**Definition 2. Decomposition:** Given  $\pi \in \Pi$ , a factorization  $d$  of  $\pi$  is called a *decomposition* of  $\pi$  whenever:

- a)  $d = \sigma_0$  if  $I(\pi) = 0$ ,
- b)  $d = (d(1), d(2), \dots, d(I(\pi)))$  if  $I(\pi) > 0$ .

We denote the set of all decompositions of a permutation  $\pi$  by  $D_\pi$ . In case there are many permutations, we refer to a decomposition also as  $d^\pi$ . Note that factorizations, as well as decompositions, of a permutation are not necessarily unique, e.g., for  $R = abc$  and  $R' = cba$ , the corresponding permutation  $\pi$  can be decomposed by  $d = (\sigma_1, \sigma_2, \sigma_1)$  as well as by  $d' = (\sigma_2, \sigma_1, \sigma_2)$ . However, once a decomposition is given, then there is an *induced path*, i.e., a sequence of linear orders, starting from  $R$  and ending at  $R'$  via elementary changes.

**Definition 3. Induced path:** Given  $R, R'$  and  $\pi \in \Pi$ , let  $d = (d(1), d(2), \dots, d(k)) \in D_\pi$  be a decomposition of  $\pi$ . A vector of linear orders  $\rho_d = (\rho_d(1), \rho_d(2), \dots, \rho_d(k+1)) \in \mathcal{L}^{k+1}$  is called a *path induced by  $d$*  between  $R, R'$  whenever:

- a)  $\rho_d$  is a path between  $R$  and  $R'$ ,
- b)  $\rho_d(i+1) = d(i) \cdot \rho_d(i)$  for all  $i = 1, 2, \dots, k$ , i.e., all consecutive linear orders in the path form elementary changes in the positions induced by the decomposition  $d$  (or  $\rho_d = (R, R)$  in case  $\pi = \sigma_0$ ).

**Remark 1.** Given  $R, R' \in \mathcal{L}$ , and the corresponding permutation  $\pi$ , if  $\rho$  is a path between  $R$  and  $R'$  then it is induced by some decomposition  $d \in D_\pi$ . Note also that the decomposition  $d \in D_\pi$  which induces the path  $\rho$  is unique. See Appendix A.5 for a visualization of this relation between decompositions and paths.

**Example 1.** Consider the linear orders  $R = abc$  and  $R' = cba$ , the corresponding permutation  $\pi$ . Let  $d$  be a decomposition of  $\pi$  such that  $d = (\sigma_1, \sigma_2, \sigma_1)$ . Then the path induced by  $d$ , denoted by  $\rho_d$ , is as follows:

$$\begin{array}{ccccccc}
R = & R_0 & R_1 & R_2 & R_3 & = & R' \\
& a & b & b & c & & \\
& b & a & c & b & & \\
& c & c & a & a & & 
\end{array}$$

Note that the decomposition  $d$  is not unique, in fact it is easy to see that  $d' = (\sigma_2, \sigma_1, \sigma_2)$  is also a decomposition of  $\pi$  leading to a different path.

## 2.2 Dissimilarity-Distance functions and properties

We use the term *dissimilarity* functions to denote what is also known as “semimetrics” in the literature and *distance* functions to denote “metrics”. The only difference, as briefly explained in the introduction, is the triangular inequality condition. We do not commit ourselves to any choice between these two options. We make it explicit when a function satisfies triangular inequality by referring to it as a distance function, rather than a dissimilarity function. Formally:

**Definition 4. Dissimilarity function:** A function  $\delta : \mathcal{L} \times \mathcal{L} \rightarrow \mathbb{R}$  is a semimetric (dissimilarity function) on the set of linear orders if it satisfies the following conditions:

- a) Non-negativity:  $\delta(R, R') \geq 0$  for all  $R, R' \in \mathcal{L}$ ,
- b) Identity of indiscernibles:  $\delta(R, R') = 0$  if and only if  $R = R'$  for all  $R, R' \in \mathcal{L}$ ,
- c) Symmetry:  $\delta(R, R') = \delta(R', R)$  for all  $R, R' \in \mathcal{L}$ .

**Definition 5. Distance function:** A function  $\delta : \mathcal{L} \times \mathcal{L} \rightarrow \mathbb{R}$  is a metric (distance function) on the set of linear orders if it is a semimetric (dissimilarity function) and, in addition, satisfies the following condition:

- d) Triangular inequality:  $\delta(R, R'') \leq \delta(R, R') + \delta(R', R'')$  for all  $R, R', R'' \in \mathcal{L}$ .

From this point on, to avoid confusion we refer to the two aforementioned functions as dissimilarity and distance functions respectively instead of semimetrics and metrics. Note that a distance function is also a dissimilarity function but the converse is not necessarily true. Next, we introduce two new conditions for dissimilarity functions. First condition, *positional neutrality*, ensures that the elementary changes in the same positions are treated impartially. Hence, a function should assign the same distance to any two pairs of linear orders that form elementary changes in the same position. Therefore, the function is neutral in the position, in the sense that, as long as the swaps in alternatives happens at the same position, it remains unchanged.

**Definition 6. Positional Neutrality:** A distance function  $\delta$  satisfies positional neutrality if for all  $k < m$  and for all elementary changes  $(R, R')$  and  $(\bar{R}, \bar{R}')$  in position  $k$ :

$$\delta(R, R') = \delta(\bar{R}, \bar{R}').$$



Note that this definition is equivalent to the following: “A dissimilarity function  $\delta$  satisfies positional neutrality if for all  $k < m$  and for all  $R, \bar{R} \in \mathcal{L}$ ,  $\delta(R, \sigma_k(R)) = \delta(\bar{R}, \sigma_k(\bar{R}))$ ”. This equivalence is due to the fact that every elementary change is associated with a unique elementary permutation.

Second condition is presented in twofold. *Decomposability*, requires that the dissimilarity between two linear orders is equal to the sum of dissimilarity assigned to each elementary change on some path between these linear orders. *Strong decomposability*, however, requires that this statement holds for *all* paths between these linear orders. We present both conditions below where quantifiers in parentheses are for the strong version.

**Definition 7. (Strong) Decomposability:** A dissimilarity function  $\delta$  satisfies (strong) decomposability if for all  $R, R'$  and for (all) some path(s)  $\rho = (R_0, R_1, \dots, R_k)$  between  $R$  and  $R'$ :

$$\delta(R, R') = \sum_{i=1}^k \delta(R_{i-1}, R_i).$$

Note that strong decomposability essentially implies that all paths between two linear orders lead to same dissimilarity. We show in Section 3 that strong decomposability is too demanding and it does not leave a lot of room to define functions.

### 3 Weighted Dissimilarity-Distance Functions

As we have already mentioned in the introduction, our aim is to differentiate among disagreements according to where they occur in preferences. We do that by means of a weight distribution with non-zero entries. Given an  $(m - 1)$ -dimensional weight vector,  $\omega = (\omega_1, \omega_2, \dots, \omega_{m-1})$ , we use the entries of the vector to measure dissimilarity of adjacent preferences (elementary changes in corresponding positions). For instance,  $\omega_1$  would be the measure of dissimilarity between the preferences  $abc$  and  $bac$ , whereas  $\omega_2$  would be interpreted as the measure of disagreement between  $abc$  and  $acb$ . Therefore once a weight vector  $\omega$  is given, it can be used to measure dissimilarity between any two adjacent preferences. We propose a weight function  $g_\omega : S \cup \{\sigma_0\} \rightarrow \mathbb{R}_{++}^{m-1}$  associated with  $\omega$  to formalize this idea in the following way:

$$g_\omega(\sigma_x) = \begin{cases} \omega_x & \text{if } x > 0 \\ 0 & \text{if } x = 0 \end{cases} \quad (1)$$

Given a permutation  $\pi$  and a factorization (possibly a decomposition)  $f = (f(1), f(2), \dots, f(t))$  of  $\pi$ , we make use of the same weight function<sup>6</sup> for the factorization (or decomposition) by setting  $g(f) = \sum_{i=1}^t g(f(i))$ . This allows us to measure dissimilarity between any two preferences including those that are not adjacent as well. Next, we define the classes of *weighted dissimilarity functions* and *weighted distance functions*.

---

<sup>6</sup>Note that  $g_\omega$  is essentially defined on the elementary/identity permutations, we slightly abuse notation and use it for sequences of elementary/identity permutations for readability.

**Definition 8.** A dissimilarity function  $\delta : \mathcal{L} \times \mathcal{L} \rightarrow \mathbb{R}$  is called a *weighted dissimilarity function* if there exists a weight vector  $\omega \in \Omega$  such that for all  $R, R' \in \mathcal{L}$  and their corresponding permutation  $\pi \in \Pi$ :

$$\delta(R, R') = g_\omega(d) \text{ for some decomposition } d \in D_\pi.$$

Similarly, a *weighted distance function* is defined as above in Definition 8 which, in addition, satisfies the triangular inequality condition. In most of the forthcoming definitions, propositions, and theorems, the absence of triangular inequality condition does not affect the results. Therefore, to save space, we will only distinguish between these two weighted classes of (dissimilarity and distance) functions by referring to them in parentheses instead of repeating the propositions. Given a weight vector  $\omega \in \mathbb{R}_{++}^{m-1}$ , we denote an associated weighted (dissimilarity-distance) function by  $\delta_\omega$ .

Note that, it is almost straightforward to see that the class of weighted (dissimilarity-distance) functions is shaped by the two conditions; positional neutrality and decomposability. In fact, the class of all (dissimilarity-distance) functions that satisfy these two conditions correspond to the class of weighted (dissimilarity-distance) functions.

**Proposition 1.** *Let  $\delta$  be a (dissimilarity-distance) function.  $\delta$  satisfies positional neutrality and decomposability if and only if  $\delta$  is a weighted (dissimilarity-distance) function for some  $\omega \in \Omega$ .*

*Proof.* **(If part)**

Let  $\delta$  be a weighted (dissimilarity-distance) function with the associated weight vector  $\omega \in \Omega$ . To show positional neutrality: take any two elementary changes  $(R, R')$  and  $(\bar{R}, \bar{R}')$  in the same position, say, in  $k < m$ . The corresponding permutation for both is, then,  $\pi = \sigma_k$ . Since  $\pi$  is an elementary permutation, there is a unique decomposition, i.e.,  $D_\pi = \{\sigma_k\}$ . Therefore  $\delta(R, R') = g_\omega(\sigma_k) = \delta(\bar{R}, \bar{R}')$ . To show decomposability: take any  $R, R' \in \mathcal{L}$ . Let  $\pi$  be the corresponding permutation. By definition, there exists a decomposition  $d \in D_\pi$  such that  $\delta(R, R') = g_\omega(d)$ , i.e.,  $\delta(R, R') = g_\omega(d(1)) + g_\omega(d(2)) + \dots + g_\omega(d(l))$  where  $l = I(\pi)$ . By Remark 1, there exists  $\rho_d = (R_0, R_1, \dots, R_l)$  which is the induced path between  $R, R'$ . This implies that  $R_i = d(i) \cdot R_{i-1}$  for each  $i = 1, 2, \dots, l$ . As  $\delta(R_{i-1}, R_i) = g_\omega(d(i))$ , this means  $\sum_{i=1}^l \delta(R_{i-1}, R_i) = \sum_{i=1}^l g_\omega(d(i)) = g_\omega(d) = \delta(R, R')$ .

**(Only if part)**

Let  $\delta$  satisfy the two conditions. By definition of (dissimilarity-distance) functions and positional neutrality, for each  $i = 1, 2, \dots, m-1$  and for all elementary changes  $(R_i, R_{i+1})$  in position  $i$ , there exists  $c_i = \delta(R_i, R_{i+1}) > 0$ . Now let  $R, R' \in \mathcal{L}$  and  $\pi \in \Pi$  be their corresponding permutation. By decomposability, there exists a path  $\rho = (R_0, R_1, \dots, R_k)$  between  $R$  and  $R'$  such that  $\delta(R, R') = \sum_{i=1}^k \delta(R_{i-1}, R_i)$ . By Remark 1, there exists  $d \in D_\pi$  which induces  $\rho$ , i.e.,  $\rho = \rho_d$ . Then for all elementary changes  $(\rho_d(i), \rho_d(i+1))$  on the path  $\rho_d$ , the (dissimilarity-distance) is  $c_x$  if and only if  $d(i) = \sigma_x$ . Then letting  $\omega = c = (c_1, c_2, \dots, c_{m-1})$ , we have that  $g_\omega(d) = \sum_{i=1}^k \delta(R_{i-1}, R_i)$ .  $\square$

An immediate question arises after Proposition 1, whether it is possible to describe a similar class of functions with strong decomposability. In the proposition below we show that the strong version of decomposability does not leave much room for variation in the class of weighted (dissimilarity-distance) functions. In fact, there exists a strong decomposable weighted (dissimilarity-distance) function if and only if the weight vector is constant. Formally:

**Proposition 2.** *Let  $\delta$  be a (dissimilarity-distance) function.  $\delta$  satisfies positional neutrality and strong decomposability if and only if  $\delta$  is a weighted (dissimilarity-distance) function for some “constant” weight vector  $\omega \in \Omega$ , i.e.,  $\omega = (c, c, \dots, c)$  for some  $c \in \mathbb{R}_{++}$*

*Proof.* The if part is straightforward. To show the only if part, let  $\delta_\omega$  satisfy strong decomposability and suppose  $\omega_i \neq \omega_{i+1}$  for some  $i \in \{1, 2, \dots, m-1\}$ . Then consider some  $R, R'$  with corresponding permutation  $\pi$  such that  $d = (\sigma_i, \sigma_{i+1}, \sigma_i) \in D_\pi$  is a decomposition of  $\pi$ . It is easy to see that there exists another decomposition<sup>7</sup>  $d' \in D_\pi$  such that  $d' = (\sigma_{i+1}, \sigma_i, \sigma_{i+1})$  (See Example 1). Then by strong decomposability,  $\delta(R, R') = g_\omega(d) = g_\omega(d')$  which contradicts  $\omega_i \neq \omega_{i+1}$ .  $\square$

We also distinguish between certain classes of weight distributions. The class of weight vectors  $\bar{\Omega} \subsetneq \mathbb{R}_{++}^{m-1}$  is interpreted as the class of monotonically decreasing weight vectors, i.e., for all  $\omega \in \bar{\Omega}$  and for all  $i \leq j$ ,  $\omega_i \geq \omega_j$ . This is particularly useful for scenarios wherein a disagreement at the top of preferences is more important than a disagreement at the bottom of preferences. Correspondingly, we denote the class of monotonically increasing weight vectors by  $\underline{\Omega}$ . The set of all possible weight vectors is denoted by  $\Omega$ , i.e., all possible  $(m-1)$ -dimensional vectors of nonnegative real numbers.

In the following subsections we focus on four examples within the class of weighted dissimilarity functions. Some of them are also distance functions depending on the choice of weight distribution. We first show that the well-known Kemeny distance (Kemeny, 1959) is a weighted distance function, which is unfortunately defined only for the constant weight vector  $\omega = (1, 1, \dots, 1)$ . As mentioned in the introduction, we would like to be able to differentiate between the positions of disagreements, hence a varying weight distribution. Therefore, an immediate extension of the Kemeny distance, to that end, is provided, *the path-minimizing function* which chooses a decomposition, ex post, depending on the distribution in the weight vector. We finally introduce the Lehmer function, and the inverse Lehmer function which are based on well-defined ex ante choices of decompositions for each permutation. We also elaborate on the conditions under which these two weighted dissimilarity functions become distance functions in Section 4.

### 3.1 Kemeny distance

Kemeny (1959) introduced a distance function which can be used to model the concept of ideological distances between strict preferences, i.e., linear orders. Interestingly, the

---

<sup>7</sup>This is due to the fact that  $\Pi$  is a permutation group. In group theory, certain groups have the property that  $\sigma_i \cdot \sigma_{i+1} \cdot \sigma_i = \sigma_{i+1} \cdot \sigma_i \cdot \sigma_{i+1}$  where  $\sigma_x$  is a generator and “.” is the group operator.

very same idea has numerous applications in other disciplines such as computer sciences, information theory, group theory etc. A list of names (probably not exhaustive) for the same concept includes; the Kendall tau distance, swap distance, inversion metric, word metric, permutation swap, the Levenshtein distance, the Damerau-Levenshtein distance, the Hamming distance, so on and so forth. In fact, prior to Kemeny (1959), the use of this distance can even be traced back to Cramer (1750). Formally:

**Definition 9. (Kemeny distance)** Given  $R, R' \in \mathcal{L}$  and their corresponding permutation  $\pi$ , the Kemeny distance  $\delta^K$  between  $R, R'$  is:

$$\delta^K(R, R') = I(\pi) = |R \setminus R'|.$$

It is easy to see that the Kemeny distance is a weighted distance function for the weight vector  $\omega = (1, 1, \dots, 1)$ . It assigns a weight of 1 each elementary change. It satisfies strong decomposability since the sum of 1's assigned to each elementary change on a path induced by any decomposition equals the size of the permutation  $I(\pi)$ . Together with Proposition 2, this implies that any weighted distance function  $\delta_\omega$  which satisfies strong decomposability is a multiple of the Kemeny distance, i.e.,  $c \times \delta^K$ , where  $c = \omega_i$  for all  $i = 1, 2, \dots, m - 1$ .

**Remark 2.** *Note that, as shown in Proposition 2, imposing strong decomposability restricts the class of weighted distance functions to only those with a constant weight vector. As we have explained in the introduction, our main motivation, however, is to find distance functions that possibly assign different weights to elementary changes in different positions. Hence we use the regular decomposability condition instead of the strong version.*

### 3.2 Path-minimizing function

As previously mentioned, the path minimizing function does not choose a decomposition for each permutation ex ante. Instead, depending on the distribution of the weights in  $\omega$ , it chooses a decomposition that induces a path with elementary changes that has a minimal sum of weights. Formally:

**Definition 10.** *Given any  $R, R'$  with  $\pi$ , and any weight vector  $\omega$ , the path-minimizing function is:*

$$\delta_\omega^{PM}(R, R') = \min_{d \in D_\pi} \{g_\omega(d)\}. \quad (2)$$

Note that for each weight vector  $\omega$ , the path-minimizing function chooses a decomposition that minimizes the weight function  $g_\omega(d)$ . We call such decompositions, *minimal decompositions* with respect to  $\omega$ . By Remark 1, there exists a uniquely induced path for each of these decompositions. We denote such paths by  $\rho^{PM}$  and call them *minimal paths* between  $R$  and  $R'$  with respect to  $\omega$ . Below we remark that the minimality of these paths are, in fact, preserved between any two point within the same path.

**Remark 3.** *Given a weight vector  $\omega$ , consider  $R, R' \in \mathcal{L}$  with  $\pi$  and a minimal path  $\rho^{PM} = (R_0, R_1, R_2, \dots, R_{k-1}, R_k)$  with respect to  $\omega$  between  $R = R_0$  and  $R' = R_k$ . Then*

any portion of this minimal path is also a minimal path between its beginning and its end with respect to the same  $\omega$ , i.e., the “subpath”  $\rho = (R_i, R_{i+1}, \dots, R_{j-1}, R_j)$ , for  $i, j \in \mathbb{N}$  such that  $0 \leq i < j < k$ , is a minimal path between  $R_i$  and  $R_j$ . Otherwise  $\rho^{PM}$  is not a minimal path between  $R$  and  $R'$ , since it could have followed a different subpath between  $R_i$  to  $R_j$ .

Note that by construction, the path-minimizing function satisfies identity of indiscernibles, nonnegativity and symmetry conditions. Therefore it is a dissimilarity function and by Equation 2, it is a weighted dissimilarity function. Finally, by Proposition 1 it satisfies positional neutrality and decomposability. In Section 4.1, we show that it also satisfies the triangular inequality condition for any weight distribution. Thus, from now on we shall call it the path-minimizing distance.

### 3.3 Lehmer function

The inverse of a permutation, according to Knuth (1998) was first defined by Rothe (see Muir, 1906). By using the diagram Rothe introduced, a list of numbers (also known, now, as the Lehmer code developed by Derrick Norman Lehmer) can be obtained for each permutation (See Example 4 in the appendix). The Lehmer code essentially was used to generate all possible permutations of any number of objects. In this work, it corresponds to a particular way of decomposing permutations.

Given any  $R, R'$  and their corresponding permutation  $\pi$ , a decomposition  $d_L \in D_\pi$  is the *winners' decomposition* if it permutes  $R$  such that  $R'(1)$  is carried to the 1<sup>st</sup> position, then  $R'(2)$  is carried to the 2<sup>nd</sup> position and so forth. Iteratively  $R'$  will be achieved. We call the path induced by this decomposition the *winners' path* and denote it by  $\rho_L$ . We illustrate the winners' decomposition and the induced path below with an example. For the formal description of the winners' decomposition, see Appendix A.2.

**Example 2.** Let  $R = abcd$  and  $R' = dcab$ . Then the winners' decomposition first permutes the alternative  $d$  to the top, thereafter  $c$  and so on. The induced path will look like:

$R_0$	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$
$a$	$a$	$a$	$(d)$	$d$	$d$
$b$	$b$	$(d)$	$a$	$a$	$(c)$
$c$	$(d)$	$b$	$b$	$(c)$	$a$
$(d)$	$c$	$c$	$(c)$	$b$	$b$

The winners' decomposition, therefore, is  $d_L^\pi = (\sigma_3, \sigma_2, \sigma_1, \sigma_3, \sigma_2)$ . This decomposition is well-defined for any two linear orders and so is the path  $\rho_L = (R_0, R_1, \dots, R_5)$ .

**Definition 11.** Given any  $R, R'$  with  $\pi$  and any weight vector  $\omega$ , the Lehmer function is:

$$\delta_\omega^L(R, R') = g_\omega(d_L^\pi). \quad (3)$$

Note that by construction, the Lehmer function satisfies identity of indiscernibles and nonnegativity conditions. The symmetry condition is shown in Proposition 5 in Appendix A.2. Therefore, it is a dissimilarity function and by Equation 3, it is a weighted dissimilarity function. Finally, by Proposition 1 it satisfies positional neutrality and decomposability. Triangular inequality is satisfied if the weights are decreasing from top to bottom. Hence it is a weighted distance function for decreasing weight vectors  $\omega \in \bar{\Omega}$ . We discuss this condition further in Section 4.1.

### 3.4 Inverse Lehmer function

An immediate dual of the winners' decomposition is the *losers' decomposition*. Given  $R, R'$  and a corresponding permutation  $\pi$ , a decomposition  $d_{IL}$  is the *losers' decomposition* if it permutes  $R$  such that  $R'(m)$  is carried to the  $m^{th}$  position, then  $R'(m-1)$  is carried to the  $(m-1)^{th}$  position and so forth. Iteratively  $R'$  will be achieved. We illustrate this decomposition and the induced *losers' path* below, denoted as  $\rho_{IL}$ , via the same linear orders as in Example 2. For the formal description of the losers' decomposition, see Appendix A.3.

**Example 3.** Let  $R = abcd$  and  $R' = dcab$ . Then the losers' decomposition first permutes the alternative  $b$  to the bottom, thereafter  $a$  and so on. The induced path will look like:

$R_0$	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$
$a$	$a$	$(a)$	$c$	$(c)$	$d$
$(b)$	$c$	$c$	$(a)$	$d$	$(c)$
$c$	$(b)$	$d$	$d$	$(a)$	$a$
$d$	$d$	$(b)$	$b$	$b$	$b$

The losers' decomposition, therefore, is  $d_{IL}^\pi = (\sigma_2, \sigma_3, \sigma_1, \sigma_2, \sigma_1)$ . This decomposition is also well-defined for any two linear orders and so is the path.

**Definition 12.** Given any  $R, R'$  with  $\pi$ , and any weight vector  $\omega$ , the inverse Lehmer function is:

$$\delta_\omega^{IL}(R, R') = g_\omega(d_{IL}^\pi). \quad (4)$$

Note that by construction, the inverse Lehmer function satisfies identity of indiscernibles and nonnegativity conditions. The symmetry condition is shown in Proposition 6 in Appendix A.3. Therefore it is a dissimilarity function and by Equation 4, it is a weighted dissimilarity function. Finally by Proposition 1 it satisfies positional neutrality and decomposability. Triangular inequality is satisfied if the weights increasing from top to bottom. Hence, it is a weighted distance function for increasing weight vectors  $\omega \in \underline{\Omega}$ . We discuss this condition further in Section 4.1.

## 4 Main Results

In this section, we first show that the path-minimizing function satisfies the triangular inequality regardless of the weight distribution. Then we show that for particular weight distributions, the Lehmer function and the inverse Lehmer functions equal the path-minimizing distance. This also helps us to distinguish those weight distributions that guarantee the triangular inequality for these functions.

### 4.1 Triangular Inequality of the Path-minimizing Distance

Below, we introduce a lemma which is crucial for proving the triangular inequality of the path-minimizing distance. The lemma argues that given a permutation, if on a sequence of elementary changes (induced by a non-reduced factorization), two adjacent alternatives are swapped twice then there can be a shorter factorization for the same permutation. Although this may sound very trivial, it turns out to be a very crucial component of the proof for the triangular inequality condition.

**Lemma 1.** *Let  $R = \dots xy \dots$  with  $\text{rank}(x, R) = i$  and  $R' = \dots xy \dots$  with  $\text{rank}(x, R') = j$  such that  $i \neq j$  and let  $\pi$  be the corresponding permutation. Let  $f = (f(1), f(2), \dots, f(t))$  be a factorization of  $\pi$  such that  $f(1) = \sigma_i$  and  $f(t) = \sigma_j$ . Then  $\tilde{f} = (f(2), f(3), \dots, f(t-1))$  is also a factorization of  $\pi$ .*

*Proof.* As  $f(1) = \sigma_i$  and  $f(t) = \sigma_j$ , the induced path  $\rho_f = (R_0, R_1, \dots, R_k, \dots, R_{t-1}, R_t)$ , with  $R_0 = R$  and  $R_t = R'$ , looks like:

$$\begin{array}{rcl} R_0 & = & \dots xy \dots \\ R_1 & = & \dots yx \dots \\ \vdots & & \\ R_k & = & \dots y \dots x \dots \\ \vdots & & \\ R_{t-1} & = & \dots yx \dots \\ R_t & = & \dots xy \dots \end{array}$$

Then, for  $R_1$  and  $R_{t-1}$  with corresponding permutation  $\bar{\pi}$ , we have  $\bar{f} = (f(2), f(3), \dots, f(t-1))$  as a factorization of  $\bar{\pi}$ . Now, for each linear order  $R_l$  for  $l = 1, 2, \dots, t-1$  in the subpath  $\rho_{\bar{\pi}}$ , consider a renaming of alternatives and write  $x$  instead of  $y$  and  $y$  instead of  $x$ . Let these new linear orders be denoted as  $R_l^{xy}$  for  $l = 1, 2, \dots, t-1$ . Note that this changes neither the corresponding permutation  $\bar{\pi}$  nor the factorization  $\bar{f}$ . Hence also for  $R_1^{xy}$  and  $R_{t-1}^{xy}$ ,  $\bar{\pi}$  is the corresponding permutation and  $\bar{f}$  is a factorization of  $\bar{\pi}$ . As  $R_1^{xy} = R_0$  and  $R_{t-1}^{xy} = R_t$ , and the corresponding permutations are unique, we conclude  $\bar{\pi} \equiv \pi$ . Therefore  $\bar{f}$  is a factorization of  $\pi$ .  $\square$

Next, we show that the path-minimizing distance satisfies the triangular inequality condition for any weight distribution. In fact, we show that the path-minimizing function

is the only robust function to changes in weight distribution. To state differently,  $\delta_\omega^{PM}$  is the only weighted distance function that satisfies triangular inequality condition for any  $\omega \in \Omega$ . This also implies that the most “comprehensive” weighted generalization of the Kemeny distance within the class of weighted distances is the path-minimizing distance.

**Theorem 1.** *Given any weight vector  $\omega \in \Omega$ , a weighted distance function  $\delta_\omega$  satisfies the triangular inequality condition if and only if  $\delta_\omega(R, R') = \delta_\omega^{PM}(R, R')$  for all  $R, R' \in \mathcal{L}$ .*

*Proof. (If part)* Take any  $\omega \in \Omega$ . Let  $\delta_\omega = \delta_\omega^{PM}$ . Take any  $R, R', R''$  and let  $\pi, \pi', \pi''$  be such that  $\pi(R) = R'$ ,  $\pi'(R) = R''$ , and  $\pi''(R') = R''$ .

Let  $d_1 \in D_\pi$ ,  $d_2 \in D_{\pi''}$ , and  $d_3 \in D_{\pi'}$  be the path minimizing decompositions for respective permutations. Let  $k_1, k_2, k_3 \in \mathbb{N}$  be the size of the these decompositions. We want to show that  $\delta_\omega^{PM}(R, R'') \leq \delta_\omega^{PM}(R, R') + \delta_\omega^{PM}(R', R'')$ , i.e.,  $g_\omega(d_3) \leq g_\omega(d_1) + g_\omega(d_2)$ . Now let  $f = (d_1, d_2)$  be a sequence of elementary permutations joining  $d_1$  and  $d_2$ , consecutively. Note that  $f$  is a factorization of  $\pi'$ , i.e.,  $f(k_1 + k_2) \cdot f(k_1 + k_2 - 1) \cdot \dots \cdot f(1) \cdot R = R''$ . Furthermore  $g_\omega(f) = g_\omega(d_1) + g_\omega(d_2)$ .

**Case 1:** If  $k_1 + k_2 = k_3$  then  $f$  is in fact a decomposition of  $\pi'$  and  $R'$  is on the path  $\rho_f = (R = R_0, R_1, \dots, R_{k_1+k_2} = R'')$  induced by this decomposition, in particular  $R' = R_{k_1} = \rho_f(k_1 + 1)$ . Then, by definition of  $\delta_\omega^{PM}$ ,  $g_\omega(d_3) \leq g_\omega(f) = g_\omega(d_1) + g_\omega(d_2)$ .

**Case 2:** If  $k_1 + k_2 > k_3$  then  $f$  is a non-reduced factorization of  $\pi'$  and there exists some pair of alternatives  $xy \in R \cap R''$  which is (unnecessarily) inverted on the path  $\rho_f$  at least twice. Then there exists  $i, j$  with  $0 \leq i < i + 1 < j < j + 1 \leq k_1 + k_2$  such that:

$$\begin{aligned} R_i &= \dots xy \dots \\ R_{i+1} &= \dots yx \dots \\ &\vdots \\ R_j &= \dots yx \dots \\ R_{j+1} &= \dots xy \dots \end{aligned}$$

Lemma 1 applies and there exists a reduction of  $f$  to  $f^{xy}$  and obviously  $g_\omega(f^{xy}) < g_\omega(f)$ . Applying Lemma 1 repeatedly for all such pairs eventually leads to a reduced factorization  $f^* = f^{x_1 y_1, \dots, x_l y_l}$  which is a decomposition of  $\pi'$ . Note that  $g_\omega(f^*) < g_\omega(f)$ . Then, by definition of  $\delta_\omega^{PM}$ ,  $g_\omega(d_3) \leq g_\omega(f^*) < g_\omega(f) = g_\omega(d_1) + g_\omega(d_2)$ .

**(Only if part)**

Let  $R, R'$ , and  $\pi$ . We will show by induction on the size of  $\pi$ , i.e.,  $I(\pi)$ .

(Induction basis) For  $I(\pi) = 1$ , by positional neutrality  $\delta_\omega = \delta_\omega^{PM}$ .

(Induction hypothesis) For  $I(\pi) = k$ , assume  $\delta_\omega = \delta_\omega^{PM}$ .

(Induction step) Let  $I(\pi) = k + 1$ . Let  $d \in D_\pi$  be the path minimizing decomposition of  $\pi$  under  $\delta_\omega^{PM}$ . Let  $\rho_d = (R = R_0, R_1, \dots, R_k, R_{k+1} = R')$  be the induced path. Consider  $R_0, R_k$  and the corresponding permutation  $\bar{\pi}$ . Let  $\bar{d} \in D_{\bar{\pi}}$  be the decomposition such that  $\rho_{\bar{d}} = (R_0, R_1, \dots, R_k)$ . By induction hypothesis,  $\delta_\omega^{PM}(R_0, R_k) = \delta_\omega(R_0, R_k)$ , by positional neutrality  $\delta_\omega^{PM}(R_k, R_{k+1}) = \delta_\omega(R_k, R_{k+1})$ . By triangular inequality  $\delta_\omega(R_0, R_k) + \delta_\omega(R_k, R_{k+1}) \geq \delta_\omega(R_0, R_{k+1})$ . Then by Remark 3 this implies  $\delta_\omega(R_0, R_k) + \delta_\omega(R_k, R_{k+1}) =$



$\delta_\omega^{PM}(R, R') \geq \delta_\omega(R, R') = \delta_\omega(R_0, R_{k+1})$ . By definition of  $\delta_\omega^{PM}$ , we conclude  $\delta_\omega^{PM}(R, R') = \delta_\omega(R, R')$ . □

## 4.2 Triangular Inequality of the Lehmer and the Inverse Lehmer Functions

Now we focus on the effects of variation in the weight distribution. It turns out that for particular classes of weight distributions, some weighted dissimilarity functions are equal. Under certain weight distributions, the Lehmer distance and the inverse Lehmer distance turns out to give the same distances with the path-minimizing distance. This leads us to find out also when they do satisfy the triangular inequality. Another usefulness of the upcoming results are due to varying levels of difficulty in calculating these weighted functions. We mention this briefly in the conclusion.

First, we show that if the weight vector is monotonically decreasing then the Lehmer function and the path-minimizing distance are equal. Then, as a corollary, the Lehmer function is a weighted distance function if and only if the weight vector is monotonically decreasing.

**Proposition 3.**  $\delta_\omega^L(R, R') = \delta_\omega^{PM}(R, R')$  for all  $R, R' \in \mathcal{L}$  if and only if  $\omega \in \bar{\Omega}$ , i.e., the Lehmer distance equals the path-minimizing distance if and only if the weight vector is decreasing.

*Proof. (If part)* Let  $\omega$  be a decreasing weight vector. We will show by induction on the size of difference between  $R, R' \in \mathcal{L}$ , i.e.,  $k = |R \setminus R'|$ , that for all decreasing weight vectors  $\omega \in \bar{\Omega}$  and for all  $R, R' \in \mathcal{L}$  and their corresponding permutation  $\pi$ ,  $\delta_\omega^L(R, R') = \delta_\omega^{PM}(R, R')$ .

(Induction basis:) Take any  $R, R' \in \mathcal{L}$  with  $\pi$  such that  $|R \setminus R'| = 1$ . Then there exists a unique decomposition  $\{d\} = D_\pi$  of  $\pi$  such that  $d \in S$ . Then, by positional neutrality,  $\delta_\omega^L(R, R') = \delta_\omega^{PM}(R, R') = g_\omega(d)$ .

(Induction hypothesis:) Take any  $R, R' \in \mathcal{L}$  with  $\pi$  such that  $|R \setminus R'| = k$ . Assume  $\delta_\omega^L(R, R') = \delta_\omega^{PM}(R, R')$ .

(Induction step:) Take any  $R, R' \in \mathcal{L}$  with  $\pi$  such that  $|R \setminus R'| = k + 1$ . Let  $R$  and  $R'$  be denoted by  $R_0$  and  $R_{k+1}$ . Let  $\rho^{PM} = (R_0, R_1, \dots, R_k, R_{k+1})$  be a minimal path induced by a minimal decomposition  $d^{PM} \in D_\pi$  of the path-minimizing distance  $\delta_\omega^{PM}(R_0, R_{k+1})$ . Suppose the elementary permutation in the last switch of the decomposition  $d^{PM}$  is in the  $i^{th}$  position of  $R_k$ , i.e.,  $d^{PM}(k+1) = \sigma_i$  and  $\sigma_i(R_k) = R_{k+1}$ . Note that by induction hypothesis,  $\delta_\omega^{PM}(R_0, R_k) = \delta_\omega^L(R_0, R_k)$ . By Remark 3 and decomposability of  $\delta_\omega^{PM}$ , we have  $\delta_\omega^{PM}(R_0, R_{k+1}) = \delta_\omega^{PM}(R_0, R_k) + \omega_i = \delta_\omega^L(R_0, R_k) + \omega_i$ .

Let  $R_k = a_1 a_2 \dots a_i a_{i+1} \dots a_m$ . Then, by construction  $R_{k+1} = a_1 a_2 \dots a_{i+1} a_i \dots a_m$ . Then, let  $\bar{\pi}$  be such that  $\bar{\pi}(R_0) = R_k$ . By construction, the Lehmer codes (See Definition 13 in the appendix) for each of the permutations,  $\pi$  and  $\bar{\pi}$  are equal except for the  $i^{th}$  and  $(i+1)^{th}$  components. Namely, for all  $t \in \{1, 2, \dots, m\} \setminus \{i, i+1\}$ ,  $|L(\bar{\pi})_t| = |L(\pi)_t|$  and therefore, for the same values of  $t$ , we also have  $d_{L_t}^{\bar{\pi}} = d_{L_t}^\pi = (\sigma_{t+|L(\pi)_t|-1}, \sigma_{t+|L(\pi)_t|-2}, \dots, \sigma_t)$

which implies  $g(d_{L_t}^{\bar{\pi}}) = g(d_{L_t}^{\pi})$ . Note that  $L(\pi)_i = \{(x, y) \in R \setminus R' \mid y = R'(i)\} = R \setminus R' \cap [UC(R'(i), R) \times R'(i)]$ . Therefore  $|L(\bar{\pi})_i| = |L(\pi)_{i+1}|$  and  $|L(\bar{\pi})_{i+1}| = |L(\pi)_i| - 1$ . Consider the Lehmer distances between  $R_0, R_k$  and  $R_0, R_{k+1}$  which are as follows:

$$\delta_{\omega}^L(R_0, R_{k+1}) = g_{\omega}(d_{L_t}^{\pi}) = g_{\omega}(d_{L_1}^{\pi}) + g_{\omega}(d_{L_2}^{\pi}) + \dots + g_{\omega}(d_{L_m}^{\pi}),$$

$$\delta_{\omega}^L(R_0, R_k) = g_{\omega}(d_{L_t}^{\bar{\pi}}) = g_{\omega}(d_{L_1}^{\bar{\pi}}) + g_{\omega}(d_{L_2}^{\bar{\pi}}) + \dots + g_{\omega}(d_{L_m}^{\bar{\pi}}).$$

Expanding these two expressions and subtracting the latter from the former by inserting  $|L(\pi)_{i+1}|$  instead of  $|L(\bar{\pi})_i|$ , and  $|L(\pi)_i| - 1$  instead of  $|L(\bar{\pi})_{i+1}|$ , we end up with:

$$\delta_{\omega}^L(R_0, R_{k+1}) - \delta_{\omega}^L(R_0, R_k) = g_{\omega}(\sigma_{|L(\pi)_{i+1}|+i}).$$

This implies that  $\delta_{\omega}^L(R_0, R_{k+1}) = \delta_{\omega}^L(R_0, R_k) + \omega_{|L(\pi)_{i+1}|+i}$ . Remember that  $\delta_{\omega}^{PM}(R_0, R_{k+1}) = \delta_{\omega}^L(R_0, R_k) + \omega_i$ . Furthermore since  $\omega$  is decreasing,  $\omega_{|L(\pi)_{i+1}|+i} \leq \omega_i$ . Therefore  $\delta_{\omega}^L(R_0, R_{k+1}) \leq \delta_{\omega}^{PM}(R_0, R_{k+1})$ . Hence, by definition of the path-minimizing distance,  $\delta_{\omega}^L(R_0, R_{k+1}) = \delta_{\omega}^{PM}(R_0, R_{k+1})$ .

**(Only if part)** Let  $\delta_{\omega}^L(R, R') = \delta_{\omega}^{PM}(R, R')$  for all  $R, R' \in \mathcal{L}$ . Suppose for a contradiction  $\omega$  is not decreasing, i.e. for some  $i = 1, 2, \dots, m-1$ ,  $\omega_i < \omega_{i+1}$ . Let  $\omega_i = x$  and  $\omega_{i+1} = x + \epsilon$  for some  $x, \epsilon > 0$ . Consider  $R_1, R_2, R_3, R_4, R_5, R_6 \in \mathcal{L}$  such that  $R_1 = a_1 a_2 \dots a_i a_{i+1} a_{i+2} \dots a_m$  and all six linear orders are identically ranked except for  $a_i, a_{i+1}$  and  $a_{i+2}$ :

$$\begin{aligned} R_1 &= \dots a_i a_{i+1} a_{i+2} \dots, \\ R_2 &= \dots a_i a_{i+2} a_{i+1} \dots = \sigma_{i+1} \cdot R_1, \\ R_3 &= \dots a_{i+2} a_i a_{i+1} \dots = \sigma_i \cdot R_2, \\ R_4 &= \dots a_{i+2} a_{i+1} a_i \dots = \sigma_{i+1} \cdot R_3, \\ R_5 &= \dots a_{i+1} a_{i+2} a_i \dots = \sigma_i \cdot R_4, \\ R_6 &= \dots a_{i+1} a_i a_{i+2} \dots = \sigma_{i+1} \cdot R_5. \end{aligned}$$

Note that  $|A| = m \geq 3$ , therefore  $|\mathcal{L}| \geq 6$  and the aforementioned linear orders always exist in  $\mathcal{L}$ . Then, consider the winners' path  $\rho_L$  for the Lehmer distance  $\delta_{\omega}^L(R_1, R_4)$ :

$$\rho_L = (R_1, R_2, R_3, R_4)$$

Then the Lehmer distance is:  $\delta_{\omega}^L(R_1, R_4) = \omega_{i+1} + \omega_i + \omega_{i+1}$ . Note, however, that there exists another path between  $R_1$  and  $R_4$  (in fact, the losers' path,  $\rho_{IL} = (R_1, R_6, R_5, R_4)$  which results in a distance of  $\omega_i + \omega_{i+1} + \omega_i$  (in fact, the inverse Lehmer distance). Obviously the latter path induces a smaller distance  $3x + \epsilon$  whereas the Lehmer distance is  $3x + 2\epsilon$ . Hence  $\delta_{\omega}^L(R_1, R_4) > \delta_{\omega}^{PM}(R_1, R_4)$ , which is a contradiction.  $\square$

By Proposition 3 and Theorem 1 we have the triangular inequality result for the Lehmer distance below. Note that this also means the Lehmer function is a “weighted distance function” for decreasing weight vectors.:

**Corollary 1.** *Lehmer function satisfies triangular inequality if and only if  $\omega \in \bar{\Omega}$ , i.e.,  $\omega$  is a decreasing weight vector.*

Considering the duality between the inverse Lehmer and the Lehmer functions, it is very intuitive, after Proposition 3, that the inverse Lehmer function should also be related to the path-minimizing distance in case the weight vector is inverted. First, we show that if the weight vector is monotonically increasing then the inverse Lehmer function and the path-minimizing distance are equal. Then, as a corollary, the inverse Lehmer function is a weighted distance function if and only if the weight vector is monotonically increasing.

**Proposition 4.**  $\delta_\omega^{IL}(R, R') = \delta_\omega^{PM}(R, R')$  for all  $R, R' \in \mathcal{L}$  if and only if  $\omega \in \underline{\Omega}$ , i.e., the inverse Lehmer function equals the path-minimizing distance if and only if the weight vector is increasing.

*Proof.* (**If part**) Let  $\omega$  be a increasing weight vector. Let  $R, R' \in \mathcal{L}$  with  $\pi$ . Suppose for a contradiction,  $g_\omega(d_{IL}^\pi) > g_\omega(d_{PM}^\pi)$ . Consider the inverse weight vector  $\hat{\omega}$ , and the dual permutation  $\hat{\pi}$  of  $\pi$ , and the dual decompositions  $d_L^{\hat{\pi}}$  and  $d_{PM}^{\hat{\pi}}$  of (respectively)  $d_{IL}^\pi$  and  $d_{PM}^\pi$ . By the duality argument in Section A.4, and Corollary 3, we have that  $g_{\hat{\omega}}(d_L^{\hat{\pi}}) > g_{\hat{\omega}}(d_{PM}^{\hat{\pi}})$ . Note that  $\hat{\omega} \in \bar{\Omega}$ , i.e.,  $\hat{\omega}$  is a decreasing weight vector which is a contradiction to Proposition 3.

(**Only if part**) Similar duality argument as in Proposition 3 follows.  $\square$

By Proposition 4 and Theorem 1 we have the triangular inequality result for the inverse Lehmer function. Note that this also means the inverse Lehmer function is a weighted distance function for increasing weight vectors.:

**Corollary 2.** *Inverse Lehmer distance satisfies triangular inequality if and only if  $\omega \in \underline{\Omega}$ , i.e.,  $\omega$  is an increasing weight vector.*

## 5 Conclusion

We have described a class of distance functions over linear orders that are sensitive to the positions of elementary changes and decomposable into sums of distances between elementary changes. Note that both of these properties are essential elements of the Kemeny distance. We have shown that only the path-minimizing distance satisfies the triangular inequality condition for all possible weight vectors.

We have shown that if weights are monotonically decreasing (increasing) from the upper parts of a ranking to the lower parts monotonically, then the Lehmer distance (the inverse Lehmer distance) satisfies the triangular inequality condition and is equivalent to the path-minimizing distance.

Note that finding the path-minimizing distance is not trivial. This is equivalent to a short-path problem which requires implementation of algorithms such as the algorithm of Dijkstra (1959), or finding out all possible paths between two linear orders and calculating the sums for each elementary change on these paths to obtain the minimal one. However, our results show that if the weights have a monotonic pattern (increasing or decreasing),

there is an easy way out. By calculating the Lehmer code (or a dual code for the inverse Lehmer distance), we can immediately conclude the distance of the winners' (or losers') path to be the minimal one. Since most scenarios impose a monotonic pattern on the weights, these findings are useful.

The class of weighted distance functions may also be implementable as collective preference rules, in the same fashion as the Kemeny-Young rule which assigns outcomes (strict preferences) that minimizes the Kemeny distance to a group of individual preferences. Since we have shown that the Kemeny distance is a particular case within this class, it would be interesting to see the properties of a preference rule that uses other weighted distance functions for the minimization.

Another possible line of research is to study what other conditions, the class of weighted distances satisfies. Bogart (1973) introduced several conditions for distance functions, among them, an additivity condition. This condition requires that for any three preferences  $R_1$ ,  $R_2$ , and  $R_3$  if  $R_2$  is on some path between  $R_1$  and  $R_3$  then  $\delta(R_1, R_2) + \delta(R_2, R_3) = \delta(R_1, R_3)$ . It is straightforward to see that this condition can be satisfied by all weighted distances only if weights are constant. However, one may restrict the betweenness requirement such that  $R_2$  is required to be on a path that gives the distance between  $R_1$  and  $R_3$ , e.g., the minimal path for the path-minimizing distance or the winners' path for the Lehmer distance.

Finally, we would like to point out to some work on distances over choice functions. Consider two individual choice functions, i.e., functions that choose from each possible subsets of alternatives. Klamler (2008) discusses distances on these choice functions, and analyzes the connection of these distances with the Kemeny distance for preferences. It would be interesting to extend the results therein to see a correspondence between the class weighted distances on preferences and some class of distances on choice functions.

# A Appendix

## A.1 Notation for permutation and group theory

Given any  $R, R' \in \mathcal{L}$  with  $\pi$ , consider  $M^\pi$ , the  $m \times m$  matrix form of  $\pi$  where  $M_{ij}^\pi = 1$  if and only if  $\pi(j) = i$  and  $M_{ij}^\pi = 0$  otherwise. The matrix  $M^\pi$  has entries of 1 in the intersection of the  $i^{th}$  row and the  $j^{th}$  column since the  $j^{th}$  alternative in  $R$  is equal to the  $i^{th}$  alternative in  $R'$ . Note that this particular notation has its own advantages, e.g., when the linear orders  $R, R'$  are written as a  $m \times 1$  column vectors, we have  $R' = M^\pi \cdot R$ .

Given a permutation matrix  $M^\pi$ , let us define a  $m \times m$  diagram by replacing all  $M_{kl}^\pi$  with crosses for all  $(k, l)$  with  $k < \pi(l)$  and  $l < j$  for  $j$  such that  $\pi(j) = k$ . Note that such  $(k, l)$ 's are the indices with zeros that come before an entry of 1 in a row and also before an entry of 1 in a column. Furthermore replace all entries with 1 by some dots. The established diagram is called the Rothe diagram<sup>8</sup> and denoted by  $\Gamma(\pi)$  where each crossed index refers to an inversion that is necessary to permute  $R$  to  $R'$ .

**Example 4.** Consider the same linear orders in Example 2,  $R = abcd$  and  $R' = dcab$  and  $\pi$  such that  $\pi(1) = 3$ ,  $\pi(2) = 4$ ,  $\pi(3) = 2$ ,  $\pi(4) = 1$ :

$$M^\pi = \begin{array}{cccc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array} \quad \Gamma(\pi) = \begin{array}{cccc|c} x & x & x & \cdot & 3 \\ x & x & \cdot & 0 & 2 \\ \cdot & 0 & 0 & 0 & 0 \\ 0 & \cdot & 0 & 0 & 0 \\ \hline 2 & 2 & 1 & 0 & \end{array}$$

**Definition 13. Lehmer Partition and Inversion Partition:** For any  $k, j = 1, 2, \dots, m$ , let  $L(\pi)_k$  denote the set of crossed indices  $(k, j)$  in the  $k^{th}$  row of the Rothe diagram  $\Gamma(\pi)$ . Then we call  $L(\pi) = (L(\pi)_k)_{k=1}^m$  the *Lehmer partition*. Similary let  $IL(\pi)_j$  denote the set of crossed indices  $(k, j)$  in the  $j^{th}$  column of the Rothe diagram  $\Gamma(\pi)$ . Then we call  $IL(\pi) = (IL(\pi)_j)_{j=1}^m$  the *inversion partition*. The vector composed of cardinalities of each component of the Lehmer partition, i.e.,  $(|L(\pi)_1|, |L(\pi)_2|, \dots, |L(\pi)_m|)$ , is known as the *Lehmer code*. The vector composed of the cardinalities of each component of the inversion partition, i.e.,  $(|IL(\pi)_1|, |IL(\pi)_2|, \dots, |IL(\pi)_m|)$ , is known as the *inversion list*.

The Lehmer code for  $\pi$  in Example 4 is the column on the right hand side of the Rothe diagram, i.e.,  $(3, 2, 0, 0)$  whereas the inversion list is the row just below the Rothe Diagram, i.e.,  $(2, 2, 1, 0)$ . The interpretation of the Lehmer code is that the alternative  $R'(1)$  has to be raised 3 times, and  $R'(2)$  has to be raised 2 times to achieve  $R'$  from  $R$  by elementary changes. The interpretation of the inversion list is that the alternative  $R(1)$  has to be lowered 2 times, and  $R(2)$  has to be lowered 2 times, and  $R(3)$  has to be lowered 1 times to achieve  $R'$  from  $R$ .

Note that  $(k, j) \in L(\pi)_k$  if and only if  $(k, j) \in IL(\pi)_j$ . Furthermore, each crossed index in  $(k, j) \in \Gamma(\pi)$  corresponds to a pair in  $R \setminus R'$  that is to be inverted. In particular,

<sup>8</sup>For the application of this diagram see Muir (1906), and Knuth (1998)

the crossed index  $(k, j) \in \Gamma(\pi)$  corresponds to the pair  $(R(j), R'(k)) \in R \setminus R'$ . Therefore  $L(\pi)_k = R \setminus R' \cap \{UC(R'(k), R) \times R'(k)\} = \{(x, y) \in R \setminus R' \mid y = R'(k)\}$ , i.e., the  $k^{th}$  component of the Lehmer partition contains the pairs  $(a, b) \in R \setminus R'$  where  $b$  is the alternative in  $k^{th}$  position of  $R'$  and  $R = .a.b.$  but  $R' = .b.a..$  Similarly,  $IL(\pi)_j = R \setminus R' \cap \{R(j) \times LC(R(j), R)\}$ , i.e., the  $j^{th}$  component of the inversion partition contains the pairs  $(a, b) \in R \setminus R'$  where  $a$  is the alternative in the  $j^{th}$  position of  $R$  and  $R = .a.b.$  but  $R' = .b.a..$

**Remark 4.** Note that  $M^\pi = (M^{\tilde{\pi}})^T$ , i.e., the permutation matrices of  $\pi$  and its conjugate (the inverse matrix)  $\tilde{\pi}$  are the transpose of each other, hence for the Rothe diagrams,  $\Gamma(\pi) = (\Gamma(\tilde{\pi}))^T$ . This implies the following relation between the Lehmer partition and the inverse partition:  $(k, j) \in L(\pi)_k$  if and only if  $(j, k) \in IL(\tilde{\pi})_j$ .

## A.2 Winners' decomposition and the Lehmer distance

The winners' decomposition introduced in Section 3.3 is also known as “canonical factorization” in Garsia (2002). It is visualized with the help of a diagram in Kassel et al. (2000), a similar diagram to that of Rothe according to Muir (1906). Remember that the winners' decomposition first raises the alternative in  $R$  that should be at the top of  $R'$ . This means the inversions in  $L(\pi)_1 = \{(x, y) \in R \setminus R' \mid y = R'(1)\}$  are made beforehand. Let  $\pi_{L_1}$  denote the permutation that raise  $R'(1)$  from its position in  $R$ , i.e.,  $rank(R'(1), R)$ , to the top of  $R$ . Obviously the unique decomposition, call it  $\{d_{L_1}\} = D_{\pi_{L_1}}$ , makes only the inversions in  $L(\pi)_1$  and looks like  $d_{L_1} = (\sigma_{|L(\pi)_1|}, \sigma_{|L(\pi)_1|-1}, \dots, \sigma_2, \sigma_1)$ . Formally, for each component of the Lehmer partition  $L(\pi)_k$ , let  $\pi_{L_k}$  denote the permutation that makes the inversions in  $L(\pi)_k$ . Then consider the decomposition  $d_{L_k} \in D_{\pi_{L_k}}$  such that  $d_{L_k} = (\sigma_{k+|L(\pi)_k|-1}, \sigma_{k+|L(\pi)_k|-2}, \dots, \sigma_{k+1}, \sigma_k)$  if  $|L(\pi)_k| > 0$  and  $d_{L_k} = (\sigma_0)$  otherwise. As  $\pi = \pi_{L_m} \cdot \pi_{L_{m-1}} \cdot \dots \cdot \pi_{L_1}$ , then  $d_L = (d_{L_1}, d_{L_2}, \dots, d_{L_m}) \in D_\pi$  is a well-defined decomposition of the permutation  $\pi$ .

Note that each inversion  $(k, l)$  in the Lehmer partition  $L(\pi)_k$  is assigned an elementary change in some position via the  $k^{th}$  component of the winners' decomposition  $d_{L_k}$  depending on the number  $k$  and the amount of crosses that occur in  $\Gamma(\pi)$  on the same row before  $(k, l)$ . In particular given  $d_{L_k} = (\sigma_{k+|L(\pi)_k|-1}, \sigma_{k+|L(\pi)_k|-2}, \dots, \sigma_{k+1}, \sigma_k)$ , the first cross on the  $k^{th}$  row is assigned  $\sigma_k$ , the second cross on the  $k^{th}$  row is assigned  $\sigma_{k+1}$  and so on. Note however that this does not necessarily imply  $(k, l)$  is assigned  $\sigma_{k+l}$ . In general a crossed entry  $(k, l) \in \Gamma(\pi)$ , is inverted by an elementary permutation of  $\sigma_{k+m}$ , whenever there are  $m$  crosses in the  $k^{th}$  row of  $\Gamma(\pi)$  before  $(k, l)$ .

**Definition 14.** A decomposition  $d_L^\pi \in D_\pi$  is called the “winners' decomposition” if it has the form:  $d_L^\pi = (d_{L_1}, d_{L_2}, \dots, d_{L_m})$  for all permutations  $\pi \in \Pi \setminus \{\sigma_0\}$  and  $d_L^\pi = \sigma_0$  for  $\pi = \sigma_0$ .

By using the values in Example 4, the elementary permutations that occur in the winners' decomposition can be visualized by the Rothe diagram as follows:

$$\begin{array}{cccc|c} x & x & x & \cdot & 3 \\ x & x & \cdot & 0 & 2 \\ \cdot & 0 & 0 & 0 & 0 \\ 0 & \cdot & 0 & 0 & 0 \\ \hline 2 & 2 & 1 & 0 & \end{array} \longrightarrow \begin{array}{cccc|c} \sigma_1 & \sigma_2 & \sigma_3 & \cdot & 3 \\ \sigma_2 & \sigma_3 & \cdot & 0 & 2 \\ \cdot & 0 & 0 & 0 & 0 \\ 0 & \cdot & 0 & 0 & 0 \\ \hline 2 & 2 & 1 & 0 & \end{array}$$

**Proposition 5.**  $\delta_\omega^L$  is symmetric.

*Proof.* Take any  $R, R' \in \mathcal{L}$ . Let  $\pi$  be the corresponding permutation and  $\tilde{\pi}$  be the conjugate of  $\pi$ , i.e.  $\pi(R) = R'$  and  $\tilde{\pi}(R') = R$ . We want to show that  $\delta_\omega^L(R, R') = \delta_\omega^L(R', R)$ , i.e.,

$$g_\omega(d_L^\pi) = g_\omega(d_L^{\tilde{\pi}}).$$

As  $M^\pi = (M^{\tilde{\pi}})^T$ , the Rothe diagrams of each permutation are also the transpose of each other. Then for any crossed index in  $(k, l) \in \Gamma(\pi)$ , there exists a crossed in index  $(l, k) \in \Gamma(\tilde{\pi})$ . As these crosses refer to an inversion in respective winners' decompositions,  $d_L^\pi$  and  $d_L^{\tilde{\pi}}$ , it is sufficient to show that each of such crossed indices correspond to an elementary change in the same position. Consider now the  $k^{th}$  component of the winners' decomposition  $d_{L_k}^\pi$ , and the  $l^{th}$  component of the winners' decomposition  $d_{L_l}^{\tilde{\pi}}$  which invert respectively  $(k, l) \in \Gamma(\pi)$  and  $(l, k) \in \Gamma$ . Let  $K^\pi = \{(x, y) \mid M_{xy}^\pi = 1 \text{ and } x < k \text{ and } y < l\}$  and  $K^{\tilde{\pi}} = \{(x, y) \mid M_{xy}^{\tilde{\pi}} = 1 \text{ and } x < l \text{ and } y < k\}$ . As  $M^\pi = (M^{\tilde{\pi}})^T$ , we have  $|K^\pi| = |K^{\tilde{\pi}}|$ . Note also that for each  $(x, y) \in K^\pi$  there will be one less cross in the  $k^{th}$  row before  $(k, l) \in \Gamma(\pi)$  and one less cross in the  $l^{th}$  row before  $(l, k) \in \Gamma(\tilde{\pi})$ . Therefore  $(k, l) \in \Gamma(\pi)$  will have an elementary permutation of  $\sigma_{k+(l-1)-|K^\pi|}$ . Similarly  $(l, k) \in \Gamma(\tilde{\pi})$  will have an elementary permutation of  $\sigma_{l+(k-1)-|K^{\tilde{\pi}}|}$ . As  $k + (l - 1) - |K^\pi| = l + (k - 1) - |K^{\tilde{\pi}}|$  and the choice of  $(k, l)$  is arbitrary, this completes the proof.  $\square$

### A.3 Losers' decomposition and the Inverse Lehmer distance

Let  $r(j) = \text{rank}(R'(j), R)$  denote the position, of the  $j^{th}$  alternative of  $R'$ , in  $R$ . Remember that the losers' decomposition first lowers the alternative in  $R$  that should be at the bottom of  $R'$ . This means the inversions in  $IL(\pi)_{r(m)}$  are made beforehand. Let  $\pi_{IL_{r(m)}}$  denote the permutation that lowers  $R'(m)$  from its position in  $R$ , i.e.  $r(m) = \text{rank}(R'(m), R)$ , to the bottom of  $R$ . Obviously the unique decomposition, call it  $\{d_{IL_{r(m)}}\}$ , makes only the inversions in  $IL(\pi)_{r(m)}$  and looks like  $d_{IL_{r(m)}} = (\sigma_{r(m)}, \sigma_{r(m)+1}, \dots, \sigma_{m-2}, \sigma_{m-1})$ . Formally for each component of the inversion partition  $IL(\pi)_{r(k)}$ , let  $\pi_{IL_{r(k)}}$  denote the permutation that makes inversions in  $IL(\pi)_{r(k)}$ . Then consider the decomposition  $d_{IL_{r(k)}} \in D_{\pi_{IL_{r(k)}}}$  such that  $d_{IL_{r(k)}} = (\sigma_{k-|IL(\pi)_{r(k)}|}, \sigma_{k-|IL(\pi)_{r(k)}|+1}, \dots, \sigma_{k-1})$  if  $|IL(\pi)_{r(k)}| > 0$  and  $d_{IL_{r(k)}} = (\sigma_0)$  otherwise. As  $\pi = \pi_{IL_{r(1)}} \cdot \pi_{IL_{r(2)}} \cdot \dots \cdot \pi_{IL_{r(m-1)}} \cdot \pi_{IL_{r(m)}}$ , then  $d_{IL} = (d_{IL_{r(m)}}, d_{IL_{r(m-1)}}, \dots, d_{IL_{r(1)}}) \in D_\pi$  is a well-defined composition of the permutation  $\pi$ .

**Definition 15.** A decomposition  $d_{IL}^\pi \in D_\pi$  is called the “losers’ decomposition” if it has the form:  $d_{IL}^\pi = (d_{IL_{r(m)}}, d_{IL_{r(m-1)}}, \dots, d_{IL_{r(1)}})$  for all permutations  $\pi \in \Pi \setminus \{\sigma_0\}$  and  $d_{IL}^\pi = \sigma_0$  for  $\pi = \sigma_0$ .

**Proposition 6.**  $\delta_\omega^{IL}$  is symmetric.

*Proof.* Suppose for a contradiction  $\delta_\omega^{IL}$  is not symmetric, i.e., for some  $R, R' \in \mathcal{L}$ , we have  $\delta_\omega^{IL}(R, R') \neq \delta_\omega^{IL}(R', R)$ . Then by duality argument in Section A.4 and Corollary 3, we have that  $\delta_\omega^L(\hat{R}, \hat{R}') \neq \delta_\omega^L(\hat{R}', \hat{R})$  which contradicts the symmetry of the Lehmer distance in Proposition 5.  $\square$

## A.4 Duality Argument

Let us now dwell upon the duality between the Lehmer distance and the inverse Lehmer distance. The relationship between the two distances is not only the naming but further. In fact the winners’ decomposition of some permutations looks quite similar to another’s losers’ decomposition. Given any linear order  $R \in \mathcal{L}$ , let  $\hat{R}$  denote the inverse linear order, e.g., for  $R = abcd$ ,  $\hat{R} = dcba$ . Consider the linear orders  $R = abcd$  and  $R' = dcab$  in Example 3 and the inverse linear orders  $\hat{R}, \hat{R}'$ . Let us call the corresponding permutation of these inverse linear orders as: the *dual* of  $\pi$  and denote by  $\hat{\pi}$  (not to be confused by  $\tilde{\pi}$ , i.e., the conjugate of  $\pi$ ). Below is the losers’ path of  $\pi$  and the winners’ path of  $\hat{\pi}$ .

$R_0$	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$	$\hat{R}_0$	$\hat{R}_1$	$\hat{R}_2$	$\hat{R}_3$	$\hat{R}_4$	$\hat{R}_5$
$a$	$a$	$(\mathbf{a})$	$c$	$(\mathbf{c})$	$d$	$d$	$d$	$(b)$	$b$	$b$	$b$
$(\mathbf{b})$	$c$	$c$	$(a)$	$d$	$(c)$	$c$	$(b)$	$d$	$d$	$(a)$	$a$
$c$	$(b)$	$d$	$d$	$(a)$	$a$	$(\mathbf{b})$	$c$	$c$	$(a)$	$d$	$(c)$
$d$	$d$	$(b)$	$b$	$b$	$b$	$a$	$a$	$(\mathbf{a})$	$c$	$(\mathbf{c})$	$d$
Losers’ path $\rho_{IL}^\pi$ for $\pi$						Winners’ path $\rho_L^{\hat{\pi}}$ for $\hat{\pi}$					

As observed in the figure above, for each linear order  $R_i$  in the losers’ path for  $\pi$  its inverse linear order  $\hat{R}_i$  occurs at the exact same point in the winners’ path for  $\hat{\pi}$ . This duality is observed in all decompositions of  $\pi$  and  $\hat{\pi}$ . We state this formally in a remark.

**Remark 5.** Given any  $d \in D_\pi$  there exists a dual decomposition  $\hat{d} \in D_{\hat{\pi}}$  of  $d$  such that for all  $x = 1, 2, \dots, m-1$  and for all  $i = 1, 2, \dots, I(\pi)$ ,  $\sigma_x = d(i)$  if and only if  $\sigma_{m-x} = \hat{d}(i)$ .

Next we formalize the observation that the dual of a losers’ decomposition for  $\pi$  is the winners’ decomposition in  $\hat{\pi}$ . As the dual permutation of  $\hat{\pi}$  will be  $\hat{\hat{\pi}} = \pi$ , i.e., inverting linear orders twice will result in the original linear order, we can also conclude that the dual of a winners’ decomposition for  $\pi$ , is the losers’ decomposition for  $\hat{\pi}$ .

**Proposition 7.** Given  $R, R' \in \mathcal{L}$ ,  $\pi \in \Pi$ , and the corresponding permutation  $\hat{\pi} \in \Pi$  of the inverse linear orders  $\hat{R}, \hat{R}'$ , let  $d_{IL}^\pi = (\sigma_{a_1}, \sigma_{a_2}, \dots, \sigma_{a_k})$  denote the losers’ decomposition of  $\pi \in \Pi$ . Then:

$$d_L^{\hat{\pi}} = (\sigma_{m-a_1}, \sigma_{m-a_2}, \dots, \sigma_{m-a_k}).$$



*Proof.* Note that for each  $R_i$  on the losers' path of  $\pi$ , there exists  $\hat{R}_i$  in the winners' path of  $\hat{\pi}$ . Therefore for any  $i$ , if  $R_i, R_{i+1}$  is an elementary change in position  $k$ , then  $\hat{R}_i, \hat{R}_{i+1}$  is an elementary change in position  $m - k$ , since the latter two is the inverse linear orders of the former two. Then, the relevant elementary permutation in  $d_{IL}^\pi(i + 1) = \sigma_k$  for the former, whereas the relevant elementary permutation in  $d_L^{\hat{\pi}}(i + 1) = \sigma_{m-k}$  for the latter.  $\square$

Now, given a weight vector  $\omega \in \Omega$ , let  $\hat{\omega} \in \Omega$  be such that for all  $i = 1, 2, \dots, m - 1$ ,  $\omega_i = \hat{\omega}_{m-i}$ , i.e., the vector  $\hat{\omega}$  is the “inverse weight vector” of  $\omega$ . Next, we remark about the connection between the duality in decompositions and the weight vectors.

**Remark 6.** *Given a weight vector  $\omega$ , an associated weight function  $g_\omega$ , the total sum of weights for a decomposition  $d \in D_\pi$  is equivalent to that of the dual decomposition  $\hat{d} \in D_{\hat{\pi}}$  under the inverse weight vector  $\hat{\omega}$ . Formally, given any weight vector  $\omega \in \Omega$ , and its inverse  $\hat{\omega}$ , consider a decomposition  $d \in D_\pi$  and the dual decomposition  $\hat{d} \in D_{\hat{\pi}}$ . We have the following relation:*

$$g_\omega(d) = g_{\hat{\omega}}(\hat{d}).$$

Considering the remark above, an immediate corollary to Proposition 7 is about the duality between the Lehmer and the inverse Lehmer distances:

**Corollary 3.**  $g_\omega(d_{IL}^\pi) = g_{\hat{\omega}}(d_L^{\hat{\pi}})$ , and  $g_{\hat{\omega}}(d_{IL}^{\hat{\pi}}) = g_\omega(d_L^\pi)$ .

The significance of Proposition 7 and Corollary 3 is that we can carry most of the results in the Lehmer distance to the inverse Lehmer distance by means of the corollary above. Furthermore since Lehmer code is more often used in the literature, it also enables one to easily calculate the inverse of it by using the Lehmer code for another permutation.

## A.5 A visualization for paths and decompositions

Since  $\Pi$  is a permutation group on the set of alternatives  $A$ , it is well-known that when the number of alternatives is  $m = |A|$ , a visualization of all linear orders  $\mathcal{L}$  over  $A$  is possible with an geometric object known as permutahedron, (see Santmyer, 2007), in an  $m - 1$  dimensional space. For instance the set of all linear orders over  $A = \{a, b, c\}$  can be visualized in  $\mathbb{R}_+^2$  as a hexagon:

**Example 5.** *Consider the graph in Figure 1 where each vertex corresponds to some  $R_i$  and two vertices are connected by an edge if and only if they form an elementary change.*

Note that for two linear orders  $R_1$  and  $R_3$ , and the corresthe sequence between them  $(R_1, R_2, R_3)$  is a path and induced by the decomposition  $d = (\sigma_2, \sigma_1)$  whereas the sequence  $(R_1, R_4, R_5, R_6, R_3)$  is not a path because it is induced by a non-reduced factorization  $f = (\sigma_1, \sigma_2, \sigma_1, \sigma_2)$ .

**Example 6.** *For  $A = \{a, b, c, d\}$ , a visualization of the set of all possible linear orders  $\mathcal{L}$  can be achieved by a three dimensional permutohedron, known as truncated octahedron. In Figure 2 we provide a two dimensional reduction of the truncated octahedron as a graph.*

$$\begin{aligned}
R_1 &= abc \\
R_2 &= acb \\
R_3 &= cab \\
R_4 &= bac \\
R_5 &= bca \\
R_6 &= cba
\end{aligned}$$

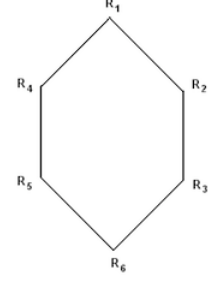


Figure 1: A graph for linear orders when  $m = 3$ .

$$\begin{array}{llllll}
R_1 = abcd & R_5 = adbc & R_9 = acdb & R_{13} = bcda & R_{17} = dbac & R_{21} = dcab \\
R_2 = abdc & R_6 = badc & R_{10} = adcb & R_{14} = cbad & R_{18} = bdca & R_{22} = dbca \\
R_3 = bacd & R_7 = bcad & R_{11} = dabc & R_{15} = cadb & R_{19} = cbda & R_{23} = cdba \\
R_4 = acbd & R_8 = cabd & R_{12} = bdac & R_{16} = dacb & R_{20} = cdab & R_{24} = dcba
\end{array}$$

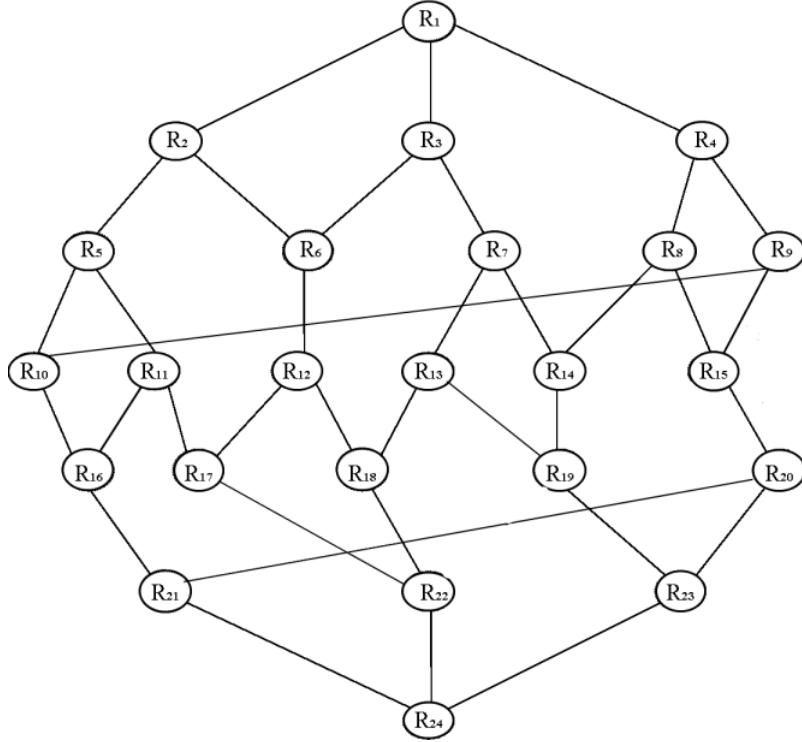


Figure 2: A graph for linear orders when  $m = 4$ .

## References

- Bogart, K. (1973). Preference structures I: Distances between transitive preference relations. *Journal of Mathematical Sociology* 3(1), 49–67.
- Cramer, G. (1750). *Introduction à l'analyse des lignes courbes algébriques*. Genève: Frères Cramer et C. Philibert.
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik* 1(1), 269–271.
- Garsia, A. (2002). The saga of reduced factorizations of elements of the symmetric group. *Publications du Laboratoire de Combinatoire et d'Informatique Mathématique* 29.
- Kassel, C., A. Lascoux, and C. Reutenauer (2000). Factorizations in Schubert cells. *Advances in Mathematics* 150(1), 1–35.
- Kemeny, J. (1959). Mathematics without numbers. *Daedalus* 88(4), 577–591.
- Klamler, C. (2008). A distance measure for choice functions. *Social Choice and Welfare* 30(3), 419–425.
- Knuth, D. (1998). *The art of computer programming: sorting and searching*, Volume 3. Reading, MA: Addison Wesley Longman Publishing Co., Inc.
- Muir, S. (1906). *The theory of determinants in the historical order of development*, Volume 1. London: Macmillan and Co., limited.